# Simulation-optimization using Agent Based Model to optimal staff allocation in Emergency Department

Antonio C. da Anunciação
*Graduate Program in Electrical Engineering*
*Universidade Federal de Minas Gerais*
Belo Horizonte, MG, Brazil
antonioemec@ufmg.br

Lucas S. Batista
*Dept. of Electrical Engineering*
*Universidade Federal de Minas Gerais*
Belo Horizonte, MG, Brazil
lusoba@ufmg.br

Michel Bessani
*Dept. of Electrical Engineering*
*Universidade Federal de Minas Gerais*
Belo Horizonte, MG, Brazil
mbessani@ufmg.br

***Abstract*—This paper presents a simheuristic approach combining agent-based modeling and multi-objective optimization for emergency department management. An agent-based simulation accurately represents patient flow and resource interactions in the emergency department, while the NSGA-II optimizes staff allocation. This hybrid method estimates Pareto-optimal solutions that balance service quality, measured as expected door-to-doctor time, and operational costs. The experiments demonstrate how different staff allocations affect the expected door-to-doctor time. The results provide actionable insights for healthcare managers, showing how simheuristics can effectively support decision-making in complex hospital environments.**

***Index Terms*—Agent-based modeling, Emergency department, Multi-objective optimization, Simulation-optimization**

## I. Introduction

The growing demand for health services, budgetary limitations, and operational inefficiencies pose significant challenges for healthcare systems worldwide. The situation is particularly critical in Brazil due to inequalities in access, shortages of professionals, and recurring overcrowding in public emergency departments (EDs) [1]. Demographic changes, epidemics, and underfunding aggravate these challenges, affecting the response capacity and sustainability of the national health system [2].

Hospitals and emergency units are often associated with long waiting times, patient dissatisfaction, and high operational stress for medical staff [3]. The motivation for this research stems from the need to improve the efficiency and quality of care in emergency units without increasing operational costs. In this context, the strategic allocation of human resources becomes crucial for ensuring timely and effective care while promoting the well-being of healthcare professionals.

Recent literature has explored various simulation and optimization techniques to enhance healthcare delivery. Agent-Based Simulation (ABS) and metaheuristic optimization methods have proven particularly effective in modeling complex systems like EDs, where agent interactions and stochastic behavior are prominent [4]. Reviews such as [5], [6] highlight the importance of integrating simulation, optimization, and operational constraints. Nevertheless, gaps still exist regarding scalable models that can dynamically adjust resource allocation in response to unpredictable demands and maintain service quality while respecting budgetary limits.

This study aims to develop an ABS of the patient flow within an ED, from arrival to discharge, incorporating the main entities, professional roles, and healthcare processes involved. The simulation model is integrated with an evolutionary optimization algorithm, specifically the NSGA-II [7], to optimize human resource allocation. This paper contributes to the field by proposing a flexible and realistic model capable of supporting healthcare managers in resource allocation decisions. The key contributions include:

- A simulation model tailored to the operational realities of Brazilian emergency units [8];
- A multiobjective optimization strategy that accounts for staffing costs, service capacity, and patient urgency.

This work builds upon the studies presented in [6], [8], [9], where an ABS model using the NetLogo framework was developed to simulate and optimize staff at Risoleta Tolentino Neves hospital in Belo Horizonte, MG, Brazil. In those studies, numerous optimization strategies and decision-making techniques are explored to reduce patient waiting times in emergency care units. Differently, this work makes simplifications due to the study timeline, data availability, and implementation methodology. The first major distinction is that the simulation model implemented here is based on *discrete-event simulation* (DES). In this approach, agents interact within a discrete time window, and the entire simulation works as a finite-state machine. In contrast, the model presented in [6], [8], [9] allows agents to move and interact continuously according to predefined behavioral rules. A finite-state machine model representing the proposed system offers a simplified yet effective way to capture key system dynamics.

Building on the simheuristic framework formalized in [10], the hybrid methodology presented in this study integrates agent-based simulation (ABS) with multi-objective optimization to address the complex challenges of emergency department (ED) staffing allocation. The approach leverages the

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), and Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG. Lucas S. Batista is a FAPEMIG-CNPQ scholarship holder (process APQ-06716-24).

complementary strengths of both paradigms: ABS captures the stochastic dynamics of patient flow and resource interactions, while optimization techniques systematically explore the trade-off space between operational costs and service quality.

The theoretical foundation of this approach draws from two key concepts in the literature. The former is the adaptive simulation strategy proposed in [10], where computational effort is dynamically allocated; initially, fewer simulations guide the metaheuristic search, and, for elite solutions, a higher quantity of replications is used to refine the better solutions. The latter is the hybrid modeling taxonomy of [11], which positions this work as a Type D Hybrid OR/MS Model, characterized by the synergistic combination of ABS with mathematical optimization techniques.

At its core, the methodology addresses a critical limitation of traditional approaches. While pure optimization struggles with EDs' stochastic patient arrivals and priority-based routing logic, standalone simulation lacks the systematic search capability to identify optimal configurations [4]. The ABS component models these operational complexities through agent interactions, generating performance metrics that feed into the optimization engine. This closed-loop integration enables the discovery of staffing solutions that balance competing objectives—minimizing patient waiting times and labor costs—while accounting for real-world uncertainties [12].

The practical implications are twofold. For ED managers, the framework provides actionable insights into capacity planning, revealing how incremental changes in staff allocation propagate through the system. For the broader field of healthcare operations, it demonstrates how hybrid modeling can bridge the gap between theoretical optimization and implementable solutions in stochastic environments. Finally, it is important to emphasize that this article does not aim to implement decision-making techniques for dynamic or real-time resource allocation.

## II. Problem Description

EDs play a critical role in healthcare systems, since they provide urgent care to patients with varying levels of severity. However, many EDs around the world are currently experiencing significant challenges, such as overcrowding, prolonged waiting times, and limited resources [13]. These challenges are often the result of increasing demand for emergency services driven by demographic changes, such as population growth and aging, as well as the emergence of new diseases.

A key issue in EDs is the excessive waiting time patients face before receiving medical attention [14]. This delay can lead to a deterioration in the patient's condition, reduced satisfaction, and higher mortality risks. One of the primary causes of these inefficiencies is the suboptimal allocation of human resources throughout the care process, encompassing reception, triage, medical consultation, and treatment. These processes involve different professional profiles (e.g., receptionists, nurses, general doctors, emergency doctors, and technicians), and poor allocation among them can result in process bottlenecks and inefficient use of available staff.

In this context, ABS (also known as agent-based modeling) [15] emerges as a powerful approach to simulate the complex and dynamic environment of EDs. Through this method, it is possible to model each participant (patients and professionals) as autonomous agents with specific behaviors, allowing the exploration of different staffing scenarios and operational policies.

This hybrid approach enables the exploration of a wide range of human resource configurations by evaluating their performance across two critical objectives: (i) minimizing the average time from patient arrival to initial medical consultation, i.e., "door-to-doctor" time, and (ii) reducing the total operational costs associated with staffing. By coupling simulation and optimization, the framework allows for identifying efficient trade-offs between quality of care and financial sustainability in emergency department management.

The system agents represent key healthcare personnel involved in the patient care process. These include doctors (both emergency specialists and general practitioners), nurses assigned to triage and treatment activities, nursing technicians responsible for clinical support tasks, and administrative assistants handling non-clinical duties within the hospital workflow.

The system is modeled using an agent-based simulation approach, wherein each healthcare professional is represented as an autonomous agent governed by a defined set of operational rules. Patients are also modeled as agents, exhibiting heterogeneous clinical profiles. Through agent interactions, the model replicates the complex dynamics of real-world hospital workflows, as illustrated in Figure 1.

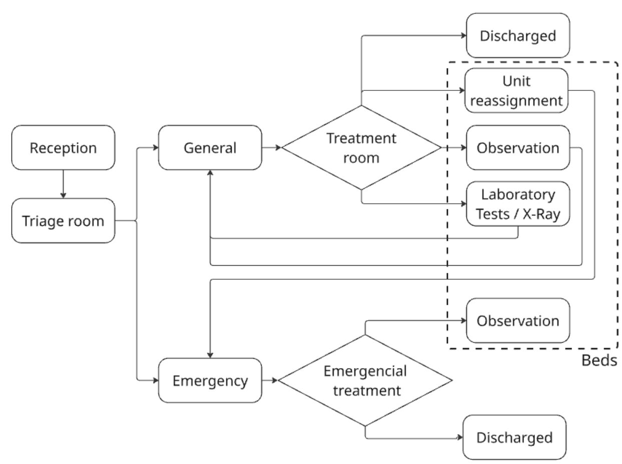


Fig. 1. Emergency Care Flowchart: Conceptual flow diagram of the emergency care system modeled using agent-based simulation. Each area represents a functional unit where agents are assigned to replicate real-world roles and resources. Patients enter through the reception and are assessed in the triage room before being directed to general or emergency care. Depending on clinical decisions, patients may be discharged, transferred to observation units, undergo laboratory testing, or be reassigned to other units. The model captures patient flow, treatment decisions, and dynamic interactions between agents to evaluate performance under stochastic conditions.

After each simulation run, the model outputs the average door-to-doctor (DTD) time. The underlying combinatorial nature of the problem introduces substantial complexity, driven by three key factors: (i) diminishing marginal returns, where staffing levels beyond the optimal threshold provide limited performance improvements; (ii) interdependencies among service areas, where alleviating one bottleneck (e.g., by adding nurses) may expose another (e.g., insufficient doctors); and (iii) physical capacity limitations inherent to the facility infrastructure.

## III. Methodology

Agent-based simulation (ABS) was adopted in this study to represent the complex and dynamic behavior of emergency department (ED) operations. Unlike aggregated models, ABS allows the explicit modeling of individual entities such as patients, nurses, and physicians, each with their own attributes and behavioral rules. This fine-grained representation enables the simulation of interactions among heterogeneous agents and the emergent system-wide behavior that results from these micro-level dynamics. For instance, Rahmat et al. [16] demonstrate how ABS can effectively capture patient flow and state deterioration under a re-triage protocol, modeling individuals as autonomous agents with evolving clinical states. Similarly, Yousefi and Ferreira [8] highlight the suitability of ABS to reflect the dynamic and adaptive environment of EDs, where agents continuously interact in a non-linear and stochastic system. These capabilities make ABS particularly appropriate for simulating healthcare environments where timing, concurrency, and human variability play crucial roles.

The simulation was developed in Python using the Mesa package [17], which provides a suitable framework for building autonomous agents, collecting data, and visualizing complex systems. All input data, parameters, and statistical distributions used in this study—whether presented in equations, tables, or simulations—are grounded in empirical observations from a real-world emergency care unit. Specifically, they are based on the operational behavior of the Risoleta Tolentino Neves Hospital, a major emergency facility located in Belo Horizonte, Minas Gerais, Brazil. These distributions were initially developed and validated by Yousefi et al. [6], [8], [9].

### A. Model Structure

The modeled environment represents a hospital divided into functional sectors, each with a defined number of beds and healthcare staff. The simulation evolves in discrete time steps, and agents interact according to predefined rules. The model includes the following agents:

- **Patient**: Represents an individual seeking medical care, with attributes such as arrival time, clinical condition, priority, waiting time, treatment duration, assigned sector, and current status. Patients follow a classification based on the Manchester Triage System (MTS) [18], which may occasionally misclassify clinical severity. Urgent cases (red or orange) are routed to emergency units, while non-urgent cases are sent to specialties like pediatrics, orthopedics, or general care.
- **Receptionist**: Registers patients sequentially and forwards them to triage after completing the registration process.
- **Triage nurse**: Performs the initial clinical assessment and assigns urgency using the MTS color-coded protocol (red, orange, yellow, green, blue) [18], directing patients to the appropriate care sector.
- **Technician nurse**: Executes diagnostic procedures, such as laboratory and radiographic tests, as requested by doctors, then refers patients back for further evaluation.
- **Treatment nurse**: Provides direct care by administering medication, performing basic procedures, and assisting doctors during treatment in clinical sectors.
- **Emergency doctor**: Delivers care to patients with urgent or life-threatening conditions, ordering diagnostics, initiating treatment, and deciding on admission, discharge, or transfer, prioritizing based on triage.
- **General doctor**: Handles non-emergency cases, conducts evaluations, prescribes treatment, and manages patient flow in general care areas, typically for triage levels yellow or green.
- **Bed**: Represents an individual unit for accommodating patients. Beds are fixed and associated with specific sectors. A patient can only be treated if a bed is available.
- **Sector**: An organizational unit that groups beds and professionals. Each sector represents areas represented in Fig. 1. The main function of this class is to record the patient flow per sector of the emergency unit.

TABLE I
Time Distributions for ED Processes.

| Process | Distribution (min) |
|---|---|
| Admission | Uniform(5,10) |
| Triage | Uniform(10,20) |
| Nursing: X-ray or Lab Test | Triangular(15,30,45) |
| Nursing: Treatment | Uniform(10,20) |
| Medical: General | Triangular(10,20,25) |
| Medical: Suturing | Triangular(15,20,40) |
| Medical: Pediatric | Triangular(10,15,30) |
| Medical: Orthopedic | Triangular(5,10,15) |
| Emergency Surgery | Triangular(10,15,30) |
| Emergency Clinical | Triangular(10,20,30) |

Table I presents the time distributions associated with each process performed by healthcare staff agents in the emergency department simulation. Each process is modeled using either a uniform or triangular probability distribution to capture its inherent variability [9]. These durations directly influence patient flow dynamics and staff workload across different medical and nursing activities, including admission, triage, diagnostics, treatment, and specialized medical care.

The proposed simulation model for patient flow in an ED is structured based on the concept of a Finite State Machine (FSM), presented in Fig. 2. In this framework, each agent in the system—such as patients, healthcare professionals, and physical resources—is modeled as an entity with a set of discrete states and transitions between them, triggered by internal or external events.

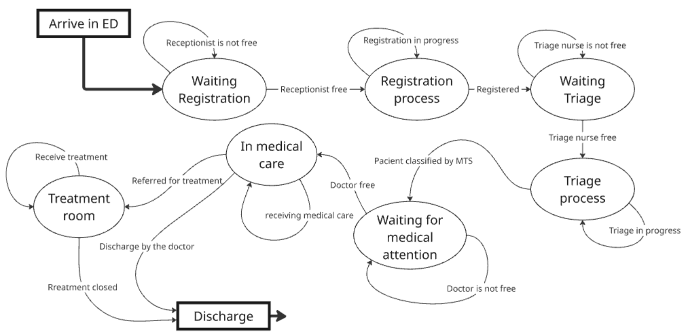


Fig. 2. Patient Finite State Machine: Representation of the patient flow within the emergency department. Each state models a phase of the patient journey, from arrival and registration through triage, waiting for medical attention, treatment, and eventual discharge. Transitions between states are governed by resource availability (e.g., admission, triage nurse, doctor) and system conditions. This state-based formalism allows for precise modeling of patient behavior and delays under stochastic and resource-constrained scenarios in the agent-based simulation framework.

The patient is the central and only mobile agent in the model, whose trajectory can be described as a sequence of well-defined states, from arrival at the emergency unit to discharge. Initially, the patient is in the arrival state, with the arrival rate according to the Fig. 3, from reference [9], transitioning to waiting for registration, then registered, undergoing triage, waiting for medical care, under treatment, undergoing tests (if necessary), and finally discharged. Each state transition depends on the availability of resources (e.g., receptionists, beds, and doctors) and clinical decisions made during the care process. This process is illustrated in Fig. 2.

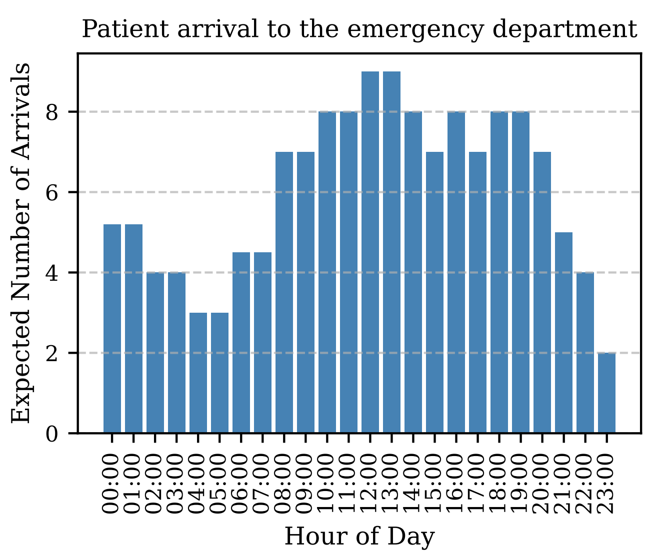


Fig. 3. Hourly distribution of expected patient arrivals at the emergency department [9]. The arrival pattern reflects typical daily fluctuations, with peaks between 11:00 and 14:00.

The agents, receptionists (admission), triage nurses, doctors, technician nurses, and beds are modeled as reactive state machines. For instance, a doctor transitions between available, evaluating a patient, awaiting test results, and finalizing treatment, while a bed shifts between available and occupied. These transitions are triggered by environmental events, such as the arrival of a new patient or the completion of a medical procedure, as illustrated in Fig. 4.

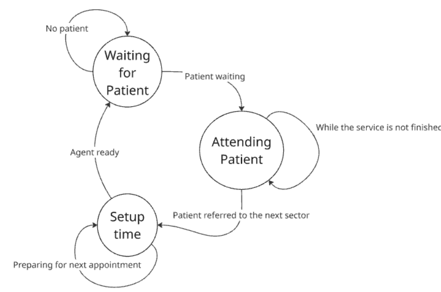


Fig. 4. Typical Agent Finite State Machine: Representing the typical behavior of a healthcare staff agent within the simulation model. The agent cycles through three main states: waiting for a patient, attending to a patient, and setup time before the next appointment. Transitions between states are triggered by the presence of patients, completion of care, and preparation tasks. This abstraction captures the agent's workload dynamics and availability, enabling the simulation to model resource contention, idle times, and service delays under stochastic patient arrivals and treatment durations.

In Fig. 4, the time that the health agent remains in the "Attending Patient" state is shown in Table I. The "Setup time" is the average time needed for a doctor to see the next patient, including both the doctor's preparation and the patient's entry into the treatment room.

Although the simulation models the entire workflow of the emergency department — including patient arrival, transfers between hospital units, staff preparation times, bed readiness, treatment durations, and medical procedures — for this study, we are solely interested in the door-to-doctor time. Therefore, any events occurring after the patient's initial encounter with a doctor are considered outside the scope of this work. This FSM-based modeling approach allows for a clear and systematic representation of the ED system's dynamics, enabling control of agent flow, monitoring of resource utilization, and identification of operational bottlenecks.

### *B. Patient Flow*

The simulated flow shown in Fig. 1 of a patient through the system includes the following steps:

1) **Arrival at the hospital**: Patients arrive dynamically based on a configurable arrival rate. Upon arrival, each patient joins a queue for registration.
2) **Registration by the receptionist**: The receptionist records patient information, assigns a unique identifier, and forwards the patient to triage. Queues may form if demand exceeds the receptionist's capacity.
3) **Triage**: A triage nurse assesses the patient's condition and assigns a priority level, which influences their placement in the treatment queue of the respective sector.

4) **Sector allocation**: The patient is assigned to an appropriate sector based on their priority. If resources (beds and staff) are available, treatment begins immediately; otherwise, the patient waits in the sector's queue.
5) **Treatment**: The patient receives care from a health professional (doctor or nurse) and occupies a bed for the duration of the treatment. Treatment time may vary depending on the severity.
6) **Discharge**: After treatment, the patient is discharged, the bed is freed, and data such as total time in the system and time in each stage are recorded.

### *C. Simulation Dynamics*

The simulation advances in discrete cycles, updating the state of each agent at every time step. The interaction rules are designed to ensure:

- A realistic, sequential patient flow;
- Efficient use of healthcare professionals and beds;
- Continuous recording of performance metrics.

The patient arrival rate to the emergency unit is presented in Fig. 3. The complete flow of a patient throughout the healthcare process is illustrated in Fig. 2. At each stage where the patient is assisted by a healthcare professional, they remain in that state for a time $t$, which follows the service time distribution specific to the corresponding professional, as detailed in Table I.

### *D. Optimization Problem Formulation*

The ED staff allocation problem can be formulated as a multi-objective integer optimization problem (1):

$$\begin{aligned} \{\mathbf{x}_1^*, \ldots, \mathbf{x}_N^*\} = \ & \arg\min_{\mathbf{x}} \quad \{\bar{t}_{\text{DTD}}(\mathbf{x}),\ C_{\text{total}}(\mathbf{x})\} \\ & \text{s.t.} \quad x_i^{\min} \le x_i \le x_i^{\max}, \quad \forall i \in \{1, 2, \ldots, k\} \\ & \qquad x_i \in \mathbb{Z}_+, \quad \forall i \in \{1, 2, \ldots, k\} \end{aligned} \tag{1}$$

where:

- $\mathbf{x}$: evaluated solution;
- $\mathbf{x}_1^*, \ldots, \mathbf{x}_N^*$: represents the set of optimal solutions;
- $\bar{t}_{DTD}$: denotes the average *door-to-doctor* time;
- $C_{\text{total}}$: represents the total daily staffing cost;
- $x_i^{\min}$, $x_i^{\max}$: lower and upper bounds of professionals allowed in category $i$, defined in Table II.

Table II represents the unit costs associated with each staff category considered in the simulation model. These values represent the per-unit cost (U.C.) of allocating one staff member of a given type, and are used in the cost component of the multi-objective optimization (2). Staff categories include administrative (e.g., Admission), nursing, technical, and medical personnel, both general and emergency-focused.

Each evaluation of (1) corresponds to a simulation run that returns the values of $t_{\text{DTD}}$. While the value of $C_{\text{total}}$ is calculated by (2):

$$C_{\text{total}}(\mathbf{x}) = \sum_{i=1}^{k} c_i x_i \tag{2}$$

where:

TABLE II
UNIT COSTS AND DOMAIN BY STAFF CATEGORY.

| **Staff member ($x_i$)** | **Unit Cost (U.C.)** | **Domain** |
|---|---|---|
| Admission | 45.00 | [1, 5] |
| Triage nurse | 75.00 | [1, 8] |
| Technician | 75.00 | [1, 6] |
| Treatment nurse | 110.00 | [1, 7] |
| Emergency doctor | 290.00 | [1, 10] |
| General doctor | 290.00 | [1, 9] |

- $c_i$ unit cost by category represented in the Table II
- $x_i$ is the number of professionals allocated to specificity $i$;
- $k$ is the number of categories (6 in the current model).

### *E. Optimization Algorithm*

To optimize the simulation-based model, we employed the Non-dominated Sorting Genetic Algorithm II (NSGA-II) [7], a well-established and robust evolutionary algorithm for solving combinatorial multi-objective problems. This choice is supported by the comparative study [19], which highlights NSGA-II as a reliable approach for low-dimensional problems with two or three objectives, particularly in discrete optimization contexts.

In our case, the optimization was conducted over two conflicting objectives, defined in (1). The first objective function, $\bar{t}_{DTD}(\mathbf{x})$, is estimated via the agent-based simulation model, which takes a staffing configuration $\mathbf{x}$ (i.e., the number of personnel allocated per role) as input and returns an estimated door-to-doctor time ($t_{DTD}$). To ensure robustness against the stochastic variation inherent to simulation, multiple replications are performed, and the average value is considered during the evaluation.

The NSGA-II algorithm was implemented with the following components:

- **Encoding**: each individual represents a candidate solution by an integer vector, with each element corresponding to a type of work team with the respective staff category, e.g, treatment, general, and admission.
- **Initialization:** The initial population is randomly generated to ensure adequate coverage of the decision space, while strictly adhering to the predefined bounds for each agent.
- **Evaluation:** Each individual is evaluated using the ABS estimation of the $\bar{t}_{DTD}(\mathbf{x})$ and the $C_{total}(\mathbf{x})$.
- **Selection:** Tournament selection is employed, preferring individuals with better ranks and greater crowding distances.
- **Variation Operators:** To explore the space of feasible staffing configurations, the evolutionary algorithm applies two main operators: *mutation* and *crossover*.
  **Mutation:** Mutation introduces variation by randomly altering one or more elements of a vector. With a fixed probability, each position may be replaced with a randomly selected value within its allowable bounds.

**Uniform Crossover:** The crossover operator generates a new individual by combining genes from two parent solutions. For each gene (i.e., vector position), the offspring independently inherits the value from parent 1 with a probability of $1/6$, or from parent 2 otherwise. For example:

$$\text{Parent 1: } [3,\ 4,\ 2,\ 5,\ 6,\ 7]$$
$$\text{Parent 2: } [4,\ 3,\ 3,\ 4,\ 7,\ 6]$$
$$\text{Offspring: } [3,\ 3,\ 2,\ 4,\ 6,\ 6]$$

This gene-wise uniform crossover promotes diversity while biasing the inheritance towards parent 2, which contributes more frequently to the offspring. Such a mechanism allows for effective recombination of building blocks from both parents, contributing to the balance between exploration and exploitation in the search process.

- **Elitism:** The current population and offspring are merged, and the best individuals (according to dominance rank and crowding distance) are selected for the next generation.

### *F. Simheuritcs*

This study adopts a simheuristic approach to address the stochastic nature of decision-making in emergency department (ED) operations. As outlined by [10], simheuristics combine metaheuristic optimization with simulation to handle stochastic objective functions that arise in real-world systems. The proposed model fits the *Type D Hybrid OR/MS Model* classification [11], integrating discrete-event simulation (DES) with metaheuristics to capture system variability while efficiently navigating the solution space. To enhance both computational efficiency and result reliability, the number of simulation replications is dynamically adjusted: fewer replications are used during early-stage exploration, while more are allocated in later stages to refine and validate promising solutions.

## IV. Results

### *A. Simulation Results*

*1) Validation Metrics and Realism of Simulation Outputs:* The validation of the model was performed by comparing its output metrics against observed patterns in EDs. The following validation steps were executed:

1) **Arrival Rate Distribution:** The hourly arrival rate distribution in our model follows empirical data closely, with variations from 5.2 patients/hour at midnight to peaks of 9 patients/hour at midday. These rates are consistent with [9] insights into patient arrival dynamics.
2) **Sector Utilization:** The model was tested for compliance with capacity limits and realistic patient flow across sectors, e.g., reception (30 patients capacity) and treatment rooms (10 patients capacity).
3) **Door-to-doctor times:** The door-to-doctor times obtained in this study are consistent with values reported in the Brazilian healthcare context. For instance, the study in [20] reports an average door-to-doctor time of approximately 52 minutes in a large public university hospital using the MTS. This is comparable to international benchmarks, where high-performing emergency departments achieve times around 25 minutes, and typical operational targets range between 25 and 60 minutes [21], [22]. Therefore, the results presented in this work are within the expected range for emergency department operations, considering both local and international standards.

Figure 5 illustrates the hourly occupancy of Admission, General, and Emergency sectors in a steady-state agent-based simulation. Occupancy in the Admission sector peaks between 15:00 and 20:00, following the surge in patient arrivals to the emergency department observed in Figure 3. The delay between arrival peaks and occupancy reflects the triage and flow processes within the system. Emergency and General treatment sectors show lower and more stable occupancy, with Emergency reacting more directly to arrival trends. These results underscore the need for adaptive staffing and capacity planning aligned with temporal demand patterns.

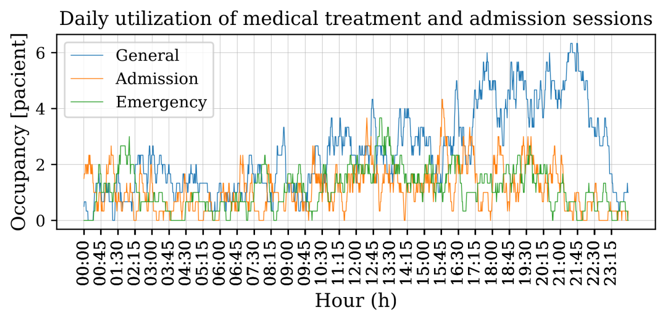


Fig. 5. One-day simulation of occupancy rates in the admission sector and medical treatment sectors (General and Emergency)

*2) Integration with Simulation-Optimization:* As described in [6], [8], [9], the integration of accurate stochastic process times is essential for resource planning in EDs. By leveraging similar distributions and extending the validation approach with empirical data, our model ensures that patient flows and resource utilizations reflect realistic ED operations. This fidelity is critical for deriving actionable insights into bottlenecks, resource constraints, and optimization opportunities.

*3) Conclusion of Validation Process:* The validation process demonstrates the robustness of our simulation framework in replicating ED workflows and timing distributions. By referencing Table I, we ensure transparency in the use of stochastic distributions, which were calibrated in alignment with the methodology of [6]. The model serves as a reliable tool for analyzing and optimizing emergency department performance.

### *B. Multi-objective Optimization Results*

The NSGA-II algorithm was executed with a population of 20 individuals and with 100 generations. A mutation rate of 10%, and a crossover rate of 16.7% for each chromosome position, a mechanism ensuring that each offspring inherits at least one gene from each parent (approximately 1/6). The number of ABS repetitions $n$ per call to evaluate $\bar{t}_{\mathrm{DTD}}$varied throughout the evolutionary process: $n = 5$ for estimations in the early generations and $n = 30$ for estimations in the final generation, as presented in [10] according to the experiment design. The model was tested on a system with an 11th Gen Intel(R) Core(TM) i5-11300H @ 3.10GHz processor and 16 GB of RAM; the total runtime to complete the optimization process was approximately 4 hours.

Figure 6 presents the trade-off between $\bar{t}_{\text{DTD}}$ and $C_{\text{total}}$. The blue circles represent all evaluated solutions, while the red "X" highlights the estimated Pareto front. As expected, solutions with lower average time tend to require more human resources, resulting in higher costs, while cost-efficient configurations typically lead to increased patient time in the system. This estimated Pareto-optimal set enables decision-makers to choose between configurations based on contextual priorities such as resource availability or service level targets.

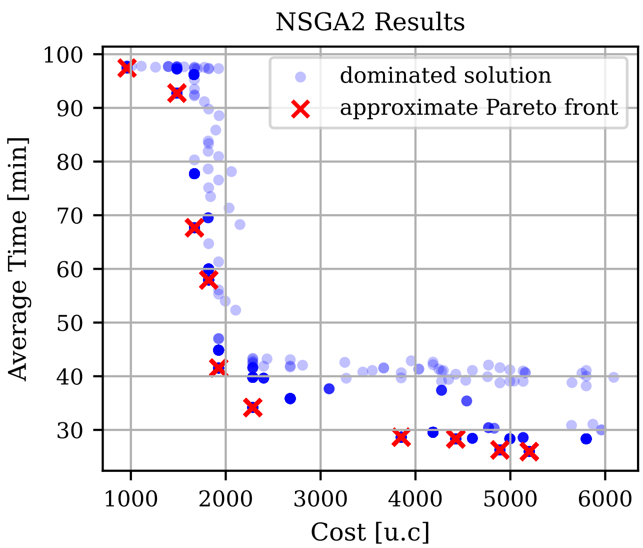


Fig. 6. Estimated Pareto front and the dominated solutions evaluated during the simheuristic approach.

### C. Analysis of Pareto-optimal Solutions

A detailed analysis of the estimated Pareto-optimal solutions, presented in Table III, reveals that certain staffing configurations yield similar average times but significantly different total costs.

This section highlights key examples that illustrate how variations in staffing decisions affect cost-effectiveness:

- **Example 1:** Solutions with nearly identical average time but different costs:
  - **Solution A:** $\bar{t}_{\text{DTD}} = 28.33$ min, $C_{\text{total}} = 4425$ u.m
  - **Solution B:** $\bar{t}_{\text{DTD}} = 28.64$ min, $C_{\text{total}} = 3850$ u.m

  The main cost reduction comes from decreasing the staff in the *Admission* area from 5 to 2. Additionally, the number of *Emergency doctors* is reduced from 7 to 6. The staffing levels for *Triage*, *Technician nurse*, *Treatment nurse*, and *General doctor* remain unchanged. Despite a marginal increase in average time (0.31 minutes, approximately 1.1%), Solution B achieves a cost reduction of 575 u.m (about 13%), indicating a significantly more cost-effective configuration.
- **Example 2:** Solutions with higher average times but substantially different costs:
  - **Solution C:** $\bar{t}_{\text{DTD}} = 92.75$ min, $C_{\text{total}} = 1485$ u.m
  - **Solution D:** $\bar{t}_{\text{DTD}} = 97.43$ min, $C_{\text{total}} = 960$ u.m

  The only staffing difference between these two solutions lies in the number of *Technician nurses*, which is reduced from 3 to 2 in Solution D. All other staffing levels remain the same. This reduction yields a significant cost saving of 525 u.m (approximately 35%), at the expense of an increase of 4.68 minutes (about 5%) in the average time.

TABLE III
ESTIMATED PARETO-OPTIMAL SOLUTIONS AND THEIR RESPECTIVE VALUES FOR $\bar{t}_{\text{DTD}}$ AND $C_{\text{TOTAL}}$ OBJECTIVE FUNCTIONS.

| **Adm.** | **Tri.** | **Tech.** | **Treat.** | **Emerg.** | **Gen.** | $\bar{t}_{\text{DTD}}$ | $C_{\text{total}}$ |
|---|---|---|---|---|---|---|---|
| 3 | 5 | 4 | 3 | 7 | 7 | 25.98 | 5200 |
| 5 | 6 | 1 | 6 | 3 | 9 | 26.27 | 4890 |
| 5 | 5 | 5 | 5 | 7 | 3 | 28.33 | 4425 |
| 2 | 5 | 3 | 5 | 6 | 3 | 28.64 | 3850 |
| 3 | 1 | 1 | 5 | 3 | 2 | 34.18 | 2285 |
| 1 | 4 | 6 | 5 | 1 | 1 | 41.53 | 1925 |
| 2 | 5 | 3 | 5 | 1 | 1 | 57.94 | 1820 |
| 2 | 3 | 3 | 5 | 1 | 1 | 67.66 | 1670 |
| 1 | 4 | 6 | 1 | 1 | 1 | 92.75 | 1485 |
| 1 | 1 | 2 | 1 | 1 | 1 | 97.43 | 960 |

These comparisons demonstrate the importance of evaluating both objectives simultaneously. Decision-makers can choose the most appropriate solution depending on whether the context prioritizes operational efficiency (low $t_{\text{DTD}}$) or cost savings.

## V. CONCLUSION

This study presents a simheuristic framework combining ABS with the NSGA-II to address staff allocation in EDs. The approach leverages the complementary strengths of both methodologies: the ABS captures the stochastic dynamics of patient flow and resource interactions, while NSGA-II explores the trade-off space between operational costs and service quality under uncertainty.

The primary contributions of this study can be summarized as follows. First, the proposed agent-based simulation (ABS) model integrates empirically calibrated time distributions (Table I) and state-machine logic (Figures 2 and 4) to realistically emulate the operational behavior of an emergency department (ED).

Second, the application of the NSGA-II algorithm over a stochastic evaluation space yields an *approximate* Pareto front (Figure 6), due to inherent variability in simulation outcomes. The resulting trade-offs reveal that configurations with comparable $\bar{t}_{\text{DTD}}$ values (e.g., 28–34 minutes) may incur cost differences of up to 13% (Table III), underscoring the relevance of multiobjective analysis in ED staffing.

Finally, the adaptive simheuristic approach—initially allocating five replications for broad exploration, and refining selected solutions with thirty replications—effectively balances computational tractability with solution robustness. This strategy is aligned with established simheuristic principles [10], where simulation is employed not only for performance estimation but also for uncertainty quantification.

The computational cost of ABS evaluations constrained the number of NSGA-II generations (100) and population size (20). While sufficient for proof-of-concept, larger-scale studies might employ parallelization, more advanced simulation strategies, or surrogate modeling to improve Pareto front approximation. The discrete-event FSM approach simplifies continuous agent behaviors.

For ED managers, this framework provides a strategy to evaluate staffing configurations under operational uncertainty. It is important to allow risk-based decision-making and analysis. There is also a quantification of the trade-offs between service quality ($\bar{t}_{\text{DTD}}$) and budgetary constraints ($C_{\text{total}}$). Finally, results highlight the importance of this kind of analysis to (re)allocating resources (e.g., reducing admission staff by 60% with minimal time impact, as in Solution B).

Future research will explore decision-making methods that are dynamically integrated with simulation. As illustrated in Fig. 3, the demand for emergency care exhibits clear temporal patterns for the problem evaluated in this work, meaning that a staffing solution optimal at one point in time may be suboptimal at another. Therefore, workforce allocation must be responsive to demand fluctuations rather than relying on static, one-size-fits-all strategies. Since both NSGA-II and ABS are stochastic, repeated runs may yield slightly different Pareto fronts. Consequently, decision-makers should evaluate the robustness of candidate solutions through post-hoc sensitivity analysis.

## References


[1] J. M. Rodrigues and A. C. Q. Barbosa, "Human resources and efficiency: a study in brazilian small hospitals," *Nova Economia*, vol. 31, pp. 217–245, 2021.

[2] L. T. Lopes and F. P. C. d. Barros, "Gestão de recursos humanos do sus na pandemia: fragilidades nas iniciativas do ministério da saúde," *Saúde em Debate*, vol. 46, pp. 277–289, 2022.

[3] C. Lima, B. Santos, D. Andrade, F. Barbosa, F. Costa, and J. Carneiro, "Fatores contextuais e individuais associados à insatisfação com o atendimento de urgência e emergência no brasil," *Cadernos de Saúde Pública*, vol. 34, no. 1, p. e00175416, 2018.

[4] K. B. Ahsan, M. Alam, D. G. Morel, and M. Karim, "Emergency department resource optimisation for improved performance: a review," *Journal of Industrial Engineering International*, vol. 15, no. Suppl 1, pp. S253–S266, 2019.

[5] E. Ouda, A. Sleptchenko, and M. C. E. Simsekler, "Comprehensive review and future research agenda on discrete-event simulation and agent-based simulation of emergency departments," *Simulation Modelling Practice and Theory*, vol. 129, p. 102823, 09 2023.

[6] M. Yousefi and M. Yousefi, "Human resource allocation in an emergency department: A metamodel-based simulation optimization," *Kybernetes*, vol. 49, no. 3, pp. 779–796, 2020.

[7] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[8] M. Yousefi and R. Ferreira, "An agent-based simulation combined with group decision-making technique for improving the performance of an emergency department," *Brazilian Journal of Medical and Biological Research*, vol. 50, 03 2017.

[9] M. Yousefi, M. Yousefi, R. P. M. Ferreira, J. H. Kim, and F. S. Fogliatto, "Chaotic genetic algorithm and adaboost ensemble metamodeling approach for optimum resource planning in emergency departments," *Artificial Intelligence in Medicine*, vol. 84, pp. 23–33, 2018. [Online]. Available: https://doi.org/10.1016/j.artmed.2017.10.002

[10] A. A. Juan, J. Faulin, S. E. Grasman, M. Rabe, and G. Figueira, "A review of simheuristics..." *Operations Research Perspectives*, vol. 2, pp. 62–72, 2015, review article.

[11] N. Mustafee, A. Harper, and B. S. Onggo, "Hybrid modelling and simulation (ms): Driving innovation in the theory and practice of ms," in *2020 Winter Simulation Conference (WSC)*, 2020, pp. 3140–3151.

[12] E. Cabrera, M. Taboada, M. L. Iglesias, F. Epelde, and E. Luque, "Optimization of healthcare emergency departments by agent-based simulation," *Procedia computer science*, vol. 4, pp. 1880–1889, 2011.

[13] Y. Liu, T. Moyaux, G. Bouleux, and V. Cheutet, "Hybrid simulation modelling of emergency departments for resource scheduling," *Journal of Simulation*, vol. 19, no. 2, pp. 215–230, 2023.

[14] A. Nyce, S. Gandhi, B. Freeze, J. Bosire, T. Ricca, E. Kupersmith, A. Mazzarelli, and J.-S. Rachoin, "Association of emergency department waiting times with patient experience in admitted and discharged patients," *Journal of Patient Experience*, vol. 8, p. 23743735211011404, 2021, pMID: 34179441. [Online]. Available: https://doi.org/10.1177/23743735211011404

[15] C. M. Macal and M. J. North, "Tutorial on agent-based modeling and simulation," in *Proceedings of the Winter Simulation Conference, 2005.* IEEE, 2005, pp. 14–pp.

[16] M. H. Rahmat, M. Annamalai, S. A. Halim, and R. Ahmad, "Agent-based modelling and simulation of emergency department re-triage," in *2013 IEEE Business Engineering and Industrial Applications Colloquium (BEIAC)*, 2013, pp. 219–224.

[17] E. ter Hoeven, J. Kwakkel, V. Hess, T. Pike, B. Wang, J. Kazil *et al.*, "Mesa 3: Agent-based modeling with python in 2025," *Journal of Open Source Software*, vol. 10, no. 107, p. 7668, 2025.

[18] Machester Triage Group, *Emergency Triage*. John Wiley Sons, Ltd, 2013, ch. The triage method, pp. 11–24.

[19] T. P. França, T. F. d. Q. Lafetá, L. G. A. Martins, and G. M. B. de Oliveira, "A comparative analysis of moeas considering two discrete optimization problems," in *2017 Brazilian Conference on Intelligent Systems (BRACIS)*, 2017, pp. 402–407.

[20] M. A. Silva, J. C. Oliveira, and A. P. Souza, "Tempos de espera para atendimento usando sistema de triagem de manchester em um hospital universitário," *Revista Mineira de Enfermagem*, vol. 21, no. 3, pp. e–2017, 2017. [Online]. Available: https://periodicos.ufmg.br/index.php/reme/article/view/49999

[21] E. Monthly, "11 benchmarks that should matter to eps," 2013, available at: https://epmonthly.com/article/11-benchmarks-that-should-matter-to-eps/ [Accessed May 2025].

[22] M. J. El Sayed, G. R. El-Eid, M. Saliba, R. Jabbour, and E. A. Hitti, "Improving emergency department door to doctor time and process reliability: A successful implementation of lean methodology," *Medicine*, vol. 94, no. 42, p. e1679, Oct 2015.